TWELP 1200 bps Vocoder

(version 5.0) offers speech quality and intelligibility that were previously achievable only at nearly double the bitrate. Below are the results of comparative tests with the MELPe 2400 bps vocoder, which operates at twice the bitrate.

For Digital HF Radio and other markets.

TWELP Technology Features. The vocoder is based on newest technology of speech coding called "Tri-Wave Excited Linear Prediction" (TWELP) that was developed by experts of DSPINI.

TWELP technology is a new class of vocoders that differs from any other LPC-based vocoders by:

- advance reliable method of pitch estimation
- pitch-synchronous analysis
- advance tri-wave model of excitation
- newest quantization schemes
- pitch-synchronous synthesis

Thanks to these unique features, TWELP technology provides significantly better speech quality than other well-known technologies—including AMBE+2, MELPe, ACELP, and others—at equivalent bit rates ranging from 300 bps to 4800 bps and beyond.

Additionally, unlike other low-bitrate vocoders (such as MELPe, for example), TWELP delivers much higher quality for non-speech signals, including sirens, background music, and similar audio.

Speech Quality. This is a comparison with the MELPe vocoder, which operates at 2400 bps and 1200 bps. The TWELP 1200 bps and MELPe 2400 bps and 1200 bps vocoders were tested using the ITU-T P.50 speech base in 20 different languages.

Note:
We have updated the speech database by minimizing inter-speech pauses to eliminate their impact on the evaluation results. Therefore, the numbers obtained from the quality measurements using this updated speech database differ from those previously obtained with the original speech database, where speech pauses were not removed.

The ITU-T P.862 tool was used to evaluate speech quality in terms of PESQ scores:

Speech Quality Comparison
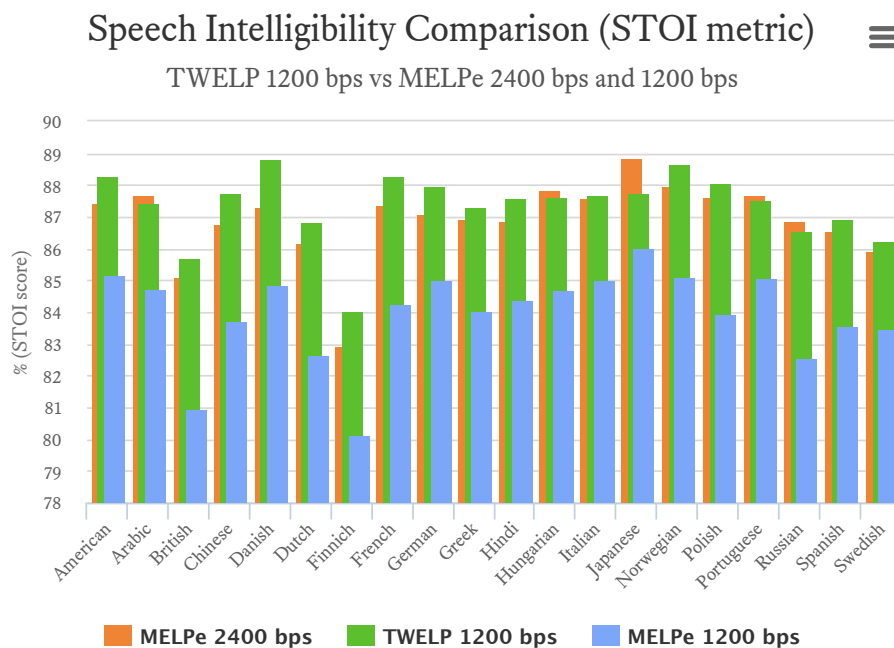### TWELP 1200 bps vs MELPe 2400 bps and 1200 bps
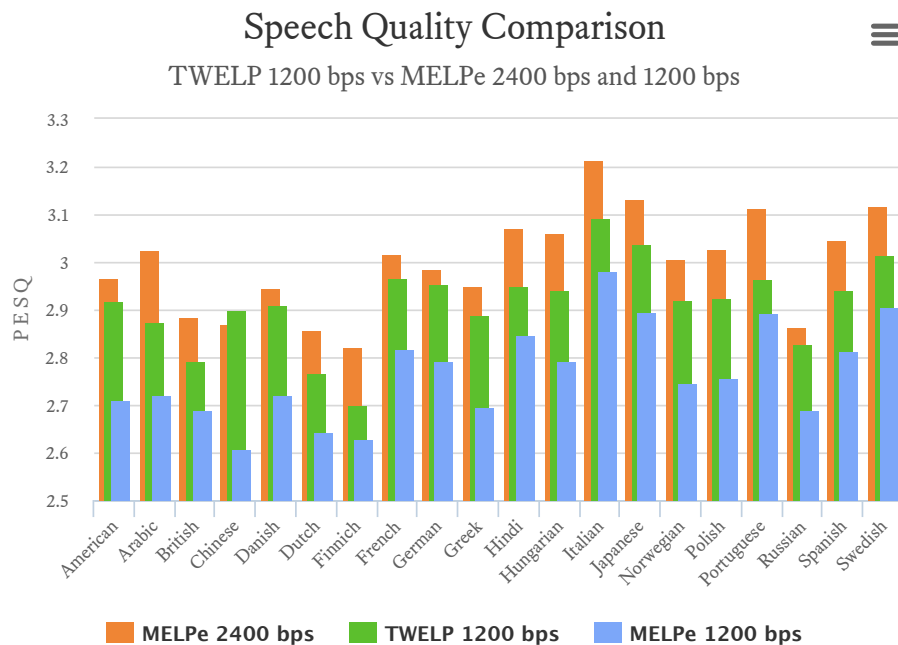
The diagram demonstrates that the speech quality of the TWELP 1200 bps vocoder is closer to that of the MELPe 2400 bps vocoder and significantly better than that of the MELPe 1200 bps vocoder. Exact numbers are shown in the table below.

Language	MELPe 2400 bps	TWELP 1200 bps	MELPe 1200 bps
American	2.967	2.917	2.711
Arabic	3.024	2.875	2.722
British	2.885	2.793	2.689
Chinese	2.871	2.899	2.607
Danish	2.946	2.910	2.720
Dutch	2.856	2.767	2.644
Finnich	2.820	2.699	2.630
French	3.018	2.969	2.818
German	2.986	2.954	2.793
Greek	2.950	2.889	2.698
Hindi	3.070	2.951	2.847
Hungarian	3.060	2.944	2.793
Italian	3.213	3.092	2.983
Japanese	3.133	3.040	2.898
Norwegian	3.007	2.920	2.745
Polish	3.027	2.926	2.758
Portuguese	3.113	2.965	2.894
Russian	2.863	2.827	2.689
Spanish	3.046	2.944	2.816
Swedish	3.119	3.013	2.907
Average	**2.999**	**2.915**	**2.768**

The average difference between MELPe 2400 and TWELP 1200 is 0.084 PESQ, while the average difference between TWELP 1200 and MELPe 1200 is 0.147 PESQ.

Speech Intelligibility. Here is a comparison with the MELPe vocoder, which operates at twice the bitrate (2400 bps) and at the same bitrate (1200 bps) . The TWELP 1200 bps vocoder and the MELPe 2400 bps and 1200 bps vocoders were tested using the ITU-T P.50 speech database, covering 20 different languages. STOI (Short-Time Objective Intelligibility) and ESTOI (Extended Short-Time Objective Intelligibility) metrics were used to assess speech intelligibility:


The diagram demonstrates the advantage of the TWELP 1200 bps vocoder over both the MELPe 1200 bps and MELPe 2400 bps vocoders, despite the latter operating at twice the bitrate. Exact values are provided in the table below:

Language	MELPe 2400 bps	TWELP 1200 bps	MELPe 1200 bps
American	87.43	88.30	85.20
Arabic	87.70	87.45	84.76
British	85.10	85.72	80.97
Chinese	86.78	87.73	83.74
Danish	87.30	88.83	84.86
Dutch	86.19	86.82	82.66
Finnich	82.94	84.05	80.12
French	87.35	88.28	84.28
German	87.10	87.97	85.01
Greek	86.96	87.30	84.03
Hindi	86.89	87.60	84.38
Hungarian	87.85	87.64	84.72
Italian	87.57	87.68	85.00
Japanese	88.88	87.75	86.01
Norwegian	87.97	88.65	85.12
Polish	87.66	88.05	83.93

Portuguese	87.68	87.56	85.05
Russian	86.89	86.56	82.55
Spanish	86.57	86.96	83.57
Swedish	85.92	86.25	83.47
Average	**86.94**	**87.36**	**83.97**

The average difference between the TWELP 1200 bps and the MELPe 2400 bps vocoders is 0.42%, while the average difference between the TWELP 1200 bps and the MELPe 1200 bps vocoders is 3.39%.

Considering that a low-bitrate vocoder is a nonlinear device that significantly distorts the spectrum of the original speech signal, the ESTOI metric provides more accurate assessments of speech intelligibility after vocoding:
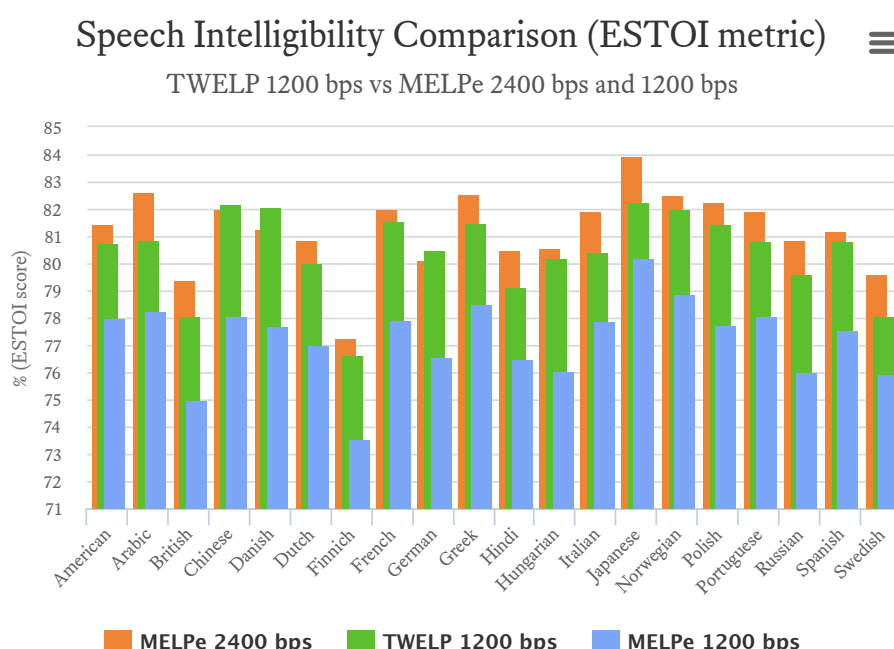
The diagram shows that speech intelligibility in the ESTOI metric after processing with the TWELP 1200 bps and MELPe 2400 bps vocoders is, on average, almost the same, despite TWELP operating at half the bitrate. In contrast, the difference between the TWELP 1200 bps and MELPe 1200 bps vocoders is significant. Exact numbers are shown in the table below.

Language	MELPe 2400 bps	TWELP 1200 bps	MELPe 1200 bps
American	81.43	80.77	78.01
Arabic	82.61	80.87	78.27
British	79.35	78.09	75.02
Chinese	82.03	82.20	78.08
Danish	81.22	82.04	77.70
Dutch	80.88	80.03	76.97
Finnich	77.26	76.62	73.56
French	82.03	81.58	77.91
German	80.14	80.52	76.58

Greek	82.55	81.49	78.51
Hindi	80.48	79.12	76.53
Hungarian	80.57	80.21	76.09
Italian	81.95	80.46	77.88
Japanese	83.96	82.28	80.21
Norwegian	82.53	81.97	78.86
Polish	82.26	81.43	77.75
Portuguese	81.94	80.80	78.04
Russian	80.86	79.61	76.02
Spanish	81.18	80.84	77.57
Swedish	79.60	78.07	75.91
Average	**81.24**	**80.45**	**76.67**

The average difference between the MELPe 2400 bps and the TWELP 1200 bps vocoders is 0.79%, while the average difference between the TWELP 1200 bps and the MELPe 1200 bps vocoders is 3.78%.

You can download the P.862 and STOI/ESTOI utilities, along with all speech samples, by using the links in the 'Downloads' section at the bottom of the page, and then check all the numbers presented above.

Speech Samples (WAV-files).
A few independent experts compared the TWELP 1200 bps vocoder with the MELPe 2400 bps and 1200 bps vocoders using the preference method.
Although opinions were almost evenly split, the majority of listeners preferred the TWELP vocoder for both MELPe bitrates, noting that the speech sound was more natural and less synthetic.

You can play and listen to short samples of the source speech, as well as the speech processed by the MELPe 2400 bps, 1200 bps vocoders and the TWELP 1200 bps vocoder, using the links in the table below.

You can also download the complete set of P.50 samples as zip files for all languages simultaneously by using the links in the 'Downloads' section at the bottom of the page.
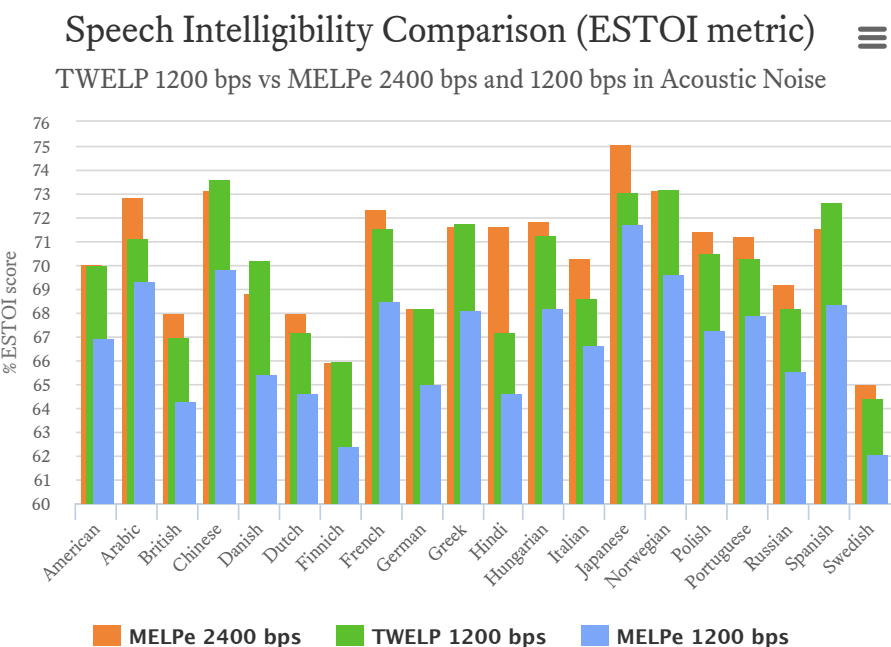
Language	Source speech	MELPe 2400 bps	MELPe 1200 bps	TWELP 1200 bps
American	▶	▶	▶	▶
Arabic	▶	▶	▶	▶
British	▶	▶	▶	▶
Chinese	▶	▶	▶	▶
Danish	▶	▶	▶	▶
Dutch	▶	▶	▶	▶
Finnich	▶	▶	▶	▶
French	▶	▶	▶	▶
German	▶	▶	▶	▶
Greek	▶	▶	▶	▶

Hindi	▶	▶	▶	▶
Hungarian	▶	▶	▶	▶
Italian	▶	▶	▶	▶
Japanese	▶	▶	▶	▶
Norwegian	▶	▶	▶	▶
Polish	▶	▶	▶	▶
Portuguese	▶	▶	▶	▶
Russian	▶	▶	▶	▶
Spanish	▶	▶	▶	▶
Swedish	▶	▶	▶	▶

Superiority In Quality Of The Non-speech Signals. In contrast to other LBR vocoders (MELPe, AMBE+2, etc.), TWELP vocoders provide high quality of non-speech signals, including police, ambulance, fire sirens, etc. This feature in conjunction with high quality natural human-sounding of voice makes TWELP vocoders well suitable for replacement of analog radio by digital radio and also for other applications where high quality transmitting of non-speech signals is relevant along with high quality transmitting of speech signals.

| Source signal | MELPe 2400 bps | MELPe 1200 bps | TWELP 1200 bps |
| ▶ | ▶ | ▶ | ▶ |

High Robustness To Acoustic Noise. In contrast to other LBR vocoders, TWELP vocoders are highly robust to acoustic noise due to a reliable pitch estimation method and other features of TWELP technology.

The diagram shows that speech intelligibility in the ESTOI metric after processing with the TWELP 1200 bps and MELPe 2400 bps vocoders is, on average, almost the same, despite TWELP operating at half the bitrate. In contrast, the difference between the TWELP 1200 bps and MELPe 1200 bps vocoders is significant. Exact numbers are shown in the table below.

| Language | MELPe 2400 bps | TWELP 1200 bps | MELPe 1200 bps |

American	70.09	69.98	66.92
Arabic	72.88	71.11	69.34
British	68.02	66.98	64.32
Chinese	73.16	73.61	69.84
Danish	68.87	70.18	65.41
Dutch	68.01	67.25	64.64
Finnich	65.96	65.97	62.45
French	72.39	71.57	68.53
German	68.23	68.20	65.00
Greek	71.61	71.77	68.12
Hindi	71.61	67.23	64.62
Hungarian	71.84	71.27	68.20
Italian	70.28	68.67	66.67
Japanese	75.04	73.08	71.71
Norwegian	73.14	73.25	69.64
Polish	71.40	70.51	67.27
Portuguese	71.25	70.31	67.93
Russian	69.19	68.18	65.57
Spanish	71.55	72.64	68.39
Swedish	65.02	64.40	62.05
Average	**70.33**	**69.81**	**65.77**

The average difference between the MELPe 2400 bps and the TWELP 1200 bps vocoders is 0.52%, while the average difference between the TWELP 1200 bps and the MELPe 1200 bps vocoders is 4.04%.

Additionally, the TWELP vocoder features an NCSE (Noise Cancellation Speech Enhancement) preprocessor, which cleans the input speech signal from noise and enhances speech quality.
Below, you can listen to a short fragment of heavily noisy English speech after passing through MELPe and TWELP vocoders, with NPP (Noise Pre-Processor) and NCSE disabled and enabled, respectively.

NPP/NCSE	Input speech (SNR=10dB)	MELPe 2400 bps	MELPe 1200 bps	TWELP 1200 bps
Disabled	▶	▶	▶	▶
Enabled	▶	▶	▶	▶

The NCSE integrated into the TWELP vocoder is described in more detail on the webpage for our standalone product, 'NCSE-AGC Preprocessor'.

High Robustness To The Channel Errors.
The TWELP technology offers highly efficient speech compression by eliminating redundancy while preserving excellent quality and intelligibility. To enhance robustness against transmission errors, we provide specialized versions called **TWELP Robust**.
These vocoders are based on an effective Joint Source-Channel Coding approach. Each vocoder is equipped

with a custom-designed FEC, tailored to its specific characteristics and operational conditions.

TWELP Robust vocoders provide high speech quality simultaneously in noisy channel as well as in noiseless channel. FEC can operate with "soft decisions" as well as with "hard decisions" from a modem. "Soft decisions" mode provides much better robustness in comparison with the "hard decisions" mode.

For all users of our non-robust vocoder versions, we offer the following recommendations.

The diagram below illustrates the sensitivity of bits at the output of the vocoder to communication channel errors.

Essentially, the diagram shows by what percentage speech quality is reduced when a specific bit is distorted. The first bits in order cause catastrophic distortions, while the latter bits have significantly less impact on quality.

Bit Sensitivity to Errors
TWELP 1200 bps vocoder

TWELP 1200 bps Vocoder
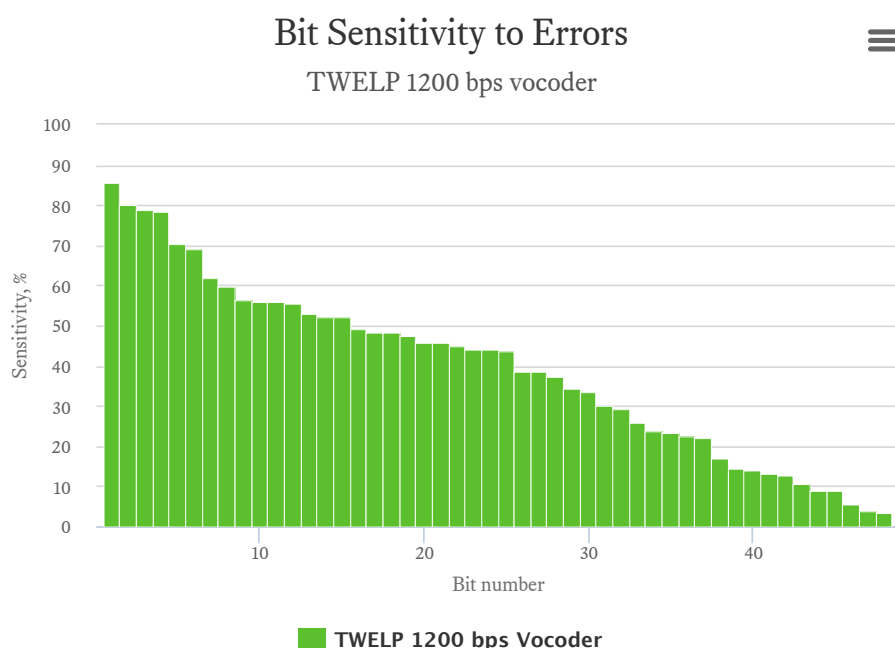
We strongly recommend using FEC (Forward Error Correction) with unequal protection of the bits in strong accordance with their sensitivity to errors and utilizing 'Soft Decisions' decoding. This will provide the highest robustness of the vocoder against errors in the channel.

 Additional Functionalities. The following additional functionalities are developed by DSPINI and integrated into TWELP vocoders:

- Noise Cancellation Speech Enhancement (NCSE)
- Automatic Gain Control (AGC),
- Voice Activity Detector (VAD),
- Discontinuous Transmission (DTX),
- Tone Detection/Generation (Single tones and Dual tones). The tones are transmitted by the vocoder facilities.

Note:
The Tone Detector/Generator functionalities are not integrated into the code by default but can be added free of charge upon request.

Each functionality has unique features, performance and characteristics, providing significant superiority over any well-known implementations on the market.

Technical Characteristics And Resource Requirements:

Technical characteristics

Bit Rate (bps)	Algorithm	Frame size (ms)	Algorithmic delay (including frame size) (ms)	Sampling rate (kHz)	Signal format	Bit stream format
1200	TWELP	40	60	8	Linear 16-bit PCM	48

Additional functionalities

Name	Functionality	Technical characteristics	
		Name	**Value**
AGC	Automatic Gain Control	Control range:	0 ... +42 dB
NCSE	Noise Canceller - Speech Enhancer	SNR increasing	20 dB
		Speech quality improvement	> 0.1 PESQ
Tone Detector	Single/Dual tones detection	In accordance with international standards	
Tone Generator	Single/Dual tones generation	Special generator, kept continuity of signal (phase and amplitude of signal of previous frame)	
DTX	Discontinuous Transmission	Reduces bit rate down to 110 bps in pauses between active speech regions	
VAD	Voice Activity Detection	High reliability even with pink noise at an SNR < 0 dB.	
CNG	Comfort Noise Generation	Type of noise	"white"
		Level	- 60 dB

The NCSE and AGC integrated into the TWELP vocoder are described in more detail on the webpage for our standalone product, 'NCSE-AGC Preprocessor'.

Resources for ARM Cortex-M4 platform

Module	MIPS* peak	Memory (KBytes)					
		Program	Data				
			Constants	Channel	Heap	Stack	
Voice Encoder	96.9	35	169	4.5	4.8	1.0	
NCSE	6.4						
AGC	0.2						
Voice Decoder	14.0						
Voice Encoder + Voice Decoder	110.9						
Total	117.5						

Resources for TI's C64 DSP platform

Module	MIPS* peak	Memory (KBytes)					
		Program	Data				
			Constants	Channel	Heap	Stack	
Voice Encoder	34.6	86	169	4.5	4.8	1.0	
NCSE	2.8						
AGC	0.1						
Voice Decoder	4.0						
Voice Encoder + Voice Decoder	38.6						
Total	41.5						

Resources (estimated) for TI's C55 DSP platform

Module	MIPS* peak	Memory (KBytes)				
		Program	Data			
			Constants	Channel	Heap	Stack
Voice Encoder	59.0	21	169	4.5	4.8	1.0
NCSE	6.7					
AGC	0.2					
Voice Decoder	10.0					
Voice Encoder + Voice Decoder	69					
Total	75.9					

* DSPINI continues optimization of the TWELP algorithm and code in order to minimize computational complexity of the vocoder.

Software Integrity and Security. DSPINI guarantees the ABSOLUTE integrity of its software, free from any undocumented features, undeclared capabilities, or hidden functions. Our customers can be assured that none of our software/code contains any secret features or functionalities concealed from the user. If necessary, we are ready to provide the source code of our software products for appropriate certification. Moreover, our software is available in source code form—you simply need to purchase the appropriate license to use it.

Guarantee And Support. DSPINI guarantees a quality and accordance of all technical characteristics of the product to requirement of current specifications. Testing and other method of quality control are used for guarantee support.

Any Platforms. DSPINI can port this vocoder software into any other DSP, RISC or general- purposes platform inshort time: 1-2 months.

Licensing Terms. To use the vocoder, customer should obtain a license from DSPINI only.

Customization. The vocoder can be customized under any specific requirements- other bit rate, frame size, any other robustness to channel errors, etc. Please contact with us for details.

Prospects. DSPINI is impoving and developing continuously a set of new vocoders with range from 300 bps up to 9600 bps, based on TWELP technology.

Related Software. This vocoder may be effectively used in a bundle with other DSPINI's products:

- Linear and acoustic echo cancellers,
- Multichannel noise cancellers (including two-microphone adaptive array),
- Wired or radiomodems for any types of channels and bitrates,

- Other products.

Downloads:

- [Datasheet (pdf)](#)
- [ITU-T P.50 source speech samples (zip)](#)
- [MELPe 2400 bps speech samples (zip)](#)
- [MELPe 1200 bps speech samples (zip)](#)
- [TWELP 1200 bps speech samples (zip)](#)
- [P.862 and STOI/ESTOI utilities](#)
- PC-evaluation package (zip) — on request
- User's Guide document (pdf) — on request